

ARTICLE OPEN



An artificial neural network for surrogate modeling of stress fields in viscoplastic polycrystalline materials

Mohammad S. Khorrami¹✉, Jaber R. Mianroodi^{1,2}✉, Nima H. Siboni^{1,2}, Pawan Goyal³, Bob Svendsen^{1,4}✉, Peter Benner³ and Dierk Raabe¹✉

The purpose of this work is the development of a trained artificial neural network for surrogate modeling of the mechanical response of elasto-viscoplastic grain microstructures. To this end, a U-Net-based convolutional neural network (CNN) is trained using results for the von Mises stress field from the numerical solution of initial-boundary-value problems (IBVPs) for mechanical equilibrium in such microstructures subject to quasi-static uniaxial extension. The resulting trained CNN (tCNN) accurately reproduces the von Mises stress field about 500 times faster than numerical solutions of the corresponding IBVP based on spectral methods. Application of the tCNN to test cases based on microstructure morphologies and boundary conditions not contained in the training dataset is also investigated and discussed.

npj Computational Materials (2023)9:37; <https://doi.org/10.1038/s41524-023-00991-z>

INTRODUCTION

In material science and engineering, material modeling is of central importance for gaining insight into the interplay between material properties, microstructure, and behavior essential for material optimization and design. In the case of polycrystalline structural materials (e.g., steels), for example, a common approach to the modeling of their mechanical behavior is the numerical solution of initial-boundary-value problems (IBVPs) for mechanical equilibrium, for example, via spectral or finite-element methods (e.g., refs. 1–7). Unfortunately, such methods are computationally quite expensive, especially at high resolution or “fidelity” (e.g., refs. 8–10). As such, alternative approaches to the modeling of complex polycrystalline and polyphase materials and their behavior under mechanical loads are the focus of current research.

Among existing alternatives to the currently used solvers, perhaps the most prominent ones are those based on artificial neural networks (ANNs) and machine learning (ML) (e.g., ref. 11). Development of corresponding models is based in particular on the fitting of ANN parameters to data via constrained optimization, i.e., ANN training, yielding a trained ANN (tANN). Except for a few early works (e.g., refs. 12,13), most such “surrogate” models based on tANNs have been introduced in the last 4 years (e.g., refs. 14–18). The data employed for training and testing can be experimental, empirical, or synthetic in nature. An example of the latter are results from the numerical solution of IBVPs based on physical models. In the “data-driven” case, training is based on such data alone. Going beyond this, one can employ physical relations (e.g., constitutive relations) on which the data are based as additional training constraints. In the case of physics-informed neural networks (PINN) (e.g., ref. 19), the data are the initial and boundary conditions of an IBVP, and the ANN is used to approximate a least-squares-based (numerical) solution of the IBVP (e.g., ref. 20). Recent reviews of applications in the field of continuum mechanics and material modeling include, for example, ref. 21. In the current work, attention is focused on the

data-driven approach. The data are obtained from the numerical solution of a BVP for quasi-static mechanical equilibrium based on viscoplastic material modeling of grains in a heterogeneous polycrystalline ensemble.

A number of data-driven approaches have been proposed for applications in solid mechanics. For example, Yang et al.²² trained a conditional generative adversarial network (cGAN) to reproduce stress and strain fields in strained isotropic elastic two-phase composites. Mianroodi et al.²³ trained a U-Net-based convolutional neural network (CNN) using results for the von Mises stress for grain microstructures consisting of isotropic elastic and ideal elastoplastic grains subject to uniaxial extension. More recently, Rashid et al.²⁴ introduced a neural-operator-based approach, the Fourier Neural Operator (FNO)^{25,26}, in particular for the surrogate modeling of stress and strain in heterogeneous composites. The U-Net-based CNN and FNO-based ANN have been compared recently by Kapoor et al.²⁷ in the context of surrogate modeling of stress fields in heterogeneous elastoplastic solids.

In the current work, a U-Net-based CNN is trained to output the von Mises stress field σ_{VM} in heterogeneous periodic microstructures consisting of inelastic grains subject to uniaxial extension. More specifically, the constitutive behavior of each grain is modeled via J_2 elasto-viscoplasticity with linear isotropic hardening. Results from the numerical solution of BVPs for quasi-static mechanical equilibrium in periodic unit cells based on this grain behavior and spectral numerical solution methods (e.g., refs. 1,2,4,7) are employed to train the CNN. For brevity, this is referred to as the “reference model” (RM) in this work. In the “Results” section, results for σ_{VM} and its average $\overline{\sigma_{VM}}$ over the unit cell U of the grain microstructure obtained from the trained CNN (tCNN) are presented and compared with corresponding results from the RM. In particular, these include the dependence of the tCNN on details of the training dataset and training such as (i) the number of grains in the microstructure, and (ii) the range of material properties chosen for each grain. In addition, the performance of the tCNN for (i) microstructure morphologies (e.g., matrix-inclusion), and (ii) extension levels, not

¹Microstructure Physics and Alloy Design, Max-Planck-Institut für Eisenforschung, Düsseldorf, Germany. ²Ergodic Labs, Lohmühlenstraße 65, 12435 Berlin, Germany.

³Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany. ⁴Material Mechanics, RWTH Aachen University, Aachen, Germany. ✉email: m.khorrami@mpie.de; j.mianroodi@mpie.de

included in the training dataset, is also investigated and discussed. The conclusions and outlook are presented in the “Discussion” section. Methods employed in this work are presented in the “Methods” section. In particular, this includes a brief summary of the model for isotropic J_2 viscoplasticity with linear isotropic constitutive response assumed for each grain in the polycrystalline ensemble. Data generation, the U-Net-based network architecture employed, and network training are also explained and discussed.

RESULTS

The purpose of this section is to compare the output of the tCNN for the von Mises stress field σ_{VM} with corresponding results from the RM for selected test cases. As explained in more detail in the “Methods” section below, the current CNN is trained with results for σ_{VM} on unit cells U with 10-grain microstructures and a resolution of 64×64 pixels subject to uniaxial extension. Training (80%) and validation (20%) datasets are based on data from 1000 such microstructures generated randomly. For simplicity, data used for training and testing are based on a single (quasi-static) extension rate of $1 \times 10^{-3} s^{-1}$. Test cases here include microstructures with (i) different numbers of grains, (ii) material property contrasts in neighboring grains, and (iii) grain morphologies, which differ from those in the training and validation datasets. The tCNN is also tested for extensions larger than those in these datasets. Lastly, the computational efficiency of the tCNN in comparison to the RM is also discussed. In what follows, σ_{RM} and σ_{tCNN} represent σ_{VM} from the RM and tCNN, respectively.

Test microstructures with a different number of grains

As stated above, the training dataset is based on 10-grain microstructures. For test cases, 64×64 pixel unit cells U with 5 and 20 grains are considered here. Each corresponding test dataset consists of 50 microstructures. The mean absolute error (MAE) for these as based on $\sigma_{tCNN} - \sigma_{RM}$ is compared with the MAE of the training dataset in Table 1. Note that the MAE increases with decreasing grain size (i.e., increasing number of grains in a unit cell of constant size) and the concomitant increase in the number of grain boundaries. As discussed in more detail below, these interfaces are regions of maximum error in σ_{tCNN} .

Example results for 5- and 20-grain test microstructures, and for a 10-grain training/validation microstructure, are shown in Figs. 1–3.

As mentioned above and shown in Figs. 1–3 (b, bottom), $|\sigma_{tCNN} - \sigma_{RM}|$ is maximal at grain boundaries and triple junctions where spatial variations in the stress field are largest and pixel resolution is lowest. In particular, the former is due to contrasts among material properties across grain boundaries. Maximum error at grain boundaries has been observed as well in the previous work (e.g.,²³). As can be seen in Figs. 1–3 (b, bottom), the spatial extent of the maximum value of the measure $|\sigma_{tCNN} - \sigma_{RM}|$ increases with increasing \bar{F}_{11} and decreasing average grain size (i.e., increased number of grains) in U .

Test microstructures with different material property distributions

Training and validation data are based on 10-grain microstructures in which material property values for the grains are selected

randomly from the ranges in Table 3. To obtain test data, random distributions of E, ν, ξ_0, h_0 have been chosen from subsets of the ranges in Table 3, i.e., $E \in [50, 300]$ GPa, $\nu \in [0.2, 0.4]$, $\xi_0 \in [50, 300]$ MPa, and $h_0 \in [0, 50]$ GPa. In particular, the four cases

$$\begin{aligned}
 \text{Case 1} \quad & E \in [50, 75] \cup [275, 300], \quad \nu \in [0.2, 0.22] \cup [0.38, 0.4], \\
 & \xi_0 \in [50, 75] \cup [275, 300], \quad h_0 \in [0, 5] \cup [45, 50], \\
 \text{Case 2} \quad & E \in [75, 100] \cup [250, 275], \quad \nu \in [0.22, 0.24] \cup [0.36, 0.38], \\
 & \xi_0 \in [75, 100] \cup [250, 275], \quad h_0 \in [5, 10] \cup [40, 45], \\
 \text{Case 3} \quad & E \in [100, 125] \cup [225, 250], \quad \nu \in [0.24, 0.26] \cup [0.34, 0.36], \\
 & \xi_0 \in [100, 125] \cup [225, 250], \quad h_0 \in [10, 15] \cup [35, 40], \\
 \text{Case 4} \quad & E \in [125, 150] \cup [200, 225], \quad \nu \in [0.26, 0.28] \cup [0.32, 0.34], \\
 & \xi_0 \in [125, 150] \cup [200, 225], \quad h_0 \in [15, 20] \cup [30, 35],
 \end{aligned} \tag{1}$$

are considered. Note the decrease in property contrast in going from Case 1 to Case 4. This is reflected in the MAEs of the corresponding test datasets shown in Table 2. Note the decrease in MAE with decreasing contrast in material properties. For example, the contrast in E is at least 215 GPa in Case 1, and 145 GPa in Case 4.

Example results for Case 1 and Case 3 are shown in Figs. 4 and 5, respectively.

Note the slight difference in the grain shape distributions in Figs. 4 and 5.

As noted above, since the material properties in each grain are homogeneous, the contrast in the property distributions is also related to the contrast in material properties at grain boundaries and triple junctions. In comparison to the results in Figs. 1–3 (b, bottom) for $|\sigma_{tCNN} - \sigma_{RM}|$ in the case that material properties values are chosen from the ranges in Table 3, note the increase in $|\sigma_{tCNN} - \sigma_{RM}|$ in Fig. 4b (bottom: up to 100 MPa) and Fig. 5b (bottom: up to 80 MPa) due to the larger contrast, especially in the former case. This is also true for the relative error shown in Figs. 4c and 5c. The figure indicates that the results from the tCNN in Case 3 (with a maximum error of 2%) are more accurate than Case 1 in which the maximum error is around 5% as shown in Fig. 4c. Training with a larger dataset in this case would result in a reduction of such errors and better agreement with the RM.

Test microstructures with different morphologies

In this subsection, the tCNN is applied to grain and microstructure morphologies not included in the training dataset. These take the form of a single inclusion embedded in a matrix material. Inclusions of circular and square form are investigated. For each, 50 test results are generated based on a random choice of material properties from the ranges in Table 3. Examples of these are shown in Figs. 6 and 7.

Analogous to the case of grain boundaries and triple junctions in the polycrystalline cases in Figs. 1–3 (b, bottom), $|\sigma_{tCNN} - \sigma_{RM}|$ is largest at the sharp matrix–inclusion (MI) interface where the contrast in material properties is greatest and pixel resolution is the lowest.

Note the maximum in $|\sigma_{tCNN} - \sigma_{RM}|$ at the MI interface in the circular case (Fig. 6b (bottom)) at $\bar{F}_{11} = 1.001$ not present at the MI interface in the square case (Fig. 7b (bottom)). This is due to the fact that (in contrast to the latter system) the former system is still elastic at $\bar{F}_{11} = 1.001$, as implied by the unit cell stress–deformation results in Fig. 6c and the larger initial resistance stress ξ_0 of the circular inclusion and corresponding contrast at the MI interface (Fig. 6a). For comparable contrasts in material properties at the (in particular sharp) interfaces of the circular and square inclusions with the matrix, one expects the largest stress concentration at the corners of the latter, and so the largest values of $|\sigma_{tCNN} - \sigma_{RM}|$ over the extension history. This is also reflected by the fact that the MAE of the test dataset for the circular inclusion case (1.42 MPa) is slightly lower than in the square inclusion case (1.58 MPa).

Table 1. MAE for test datasets based on 5- and 20-grain microstructures and for the training dataset based on 10-grain microstructures.

Number of grains	5	10	20
MAE (MPa)	1.578	1.733	1.869

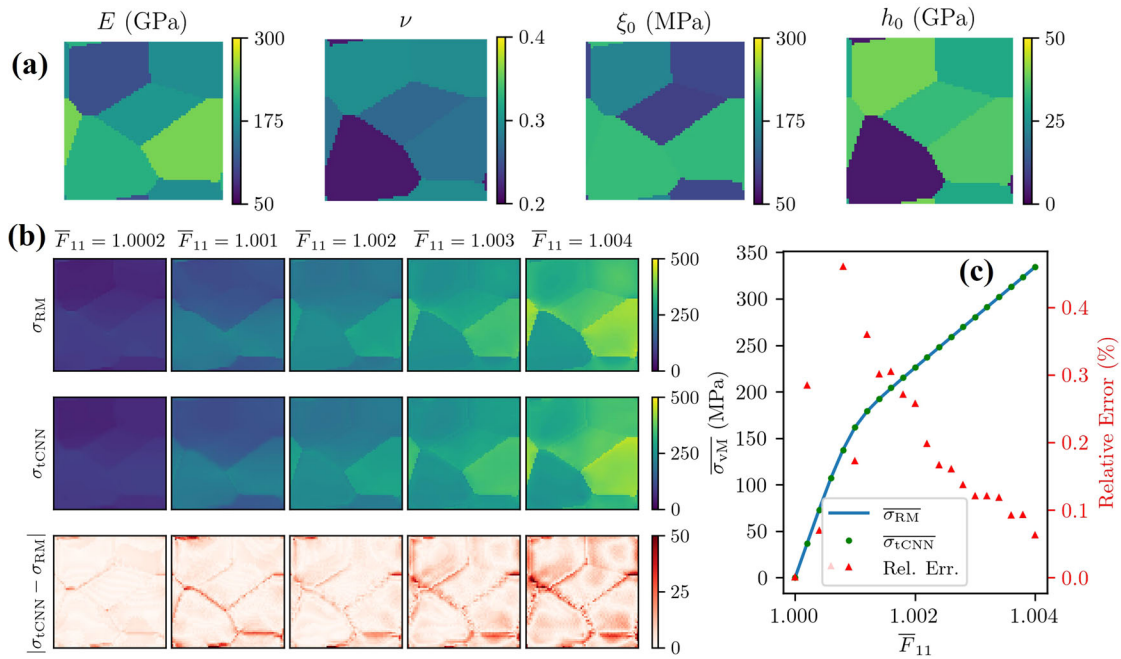


Fig. 1 Five-grain test microstructure subject to uniaxial extension along the horizontal axis. **a** Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , linear isotropic hardening modulus h_0 . **b** Results for σ_{RM} (above), σ_{tCNN} (middle), $|\sigma_{tCNN} - \sigma_{RM}|$ (below) (all MPa), at different \bar{F}_{11} . **c** $\bar{\sigma}_{RM}$ (blue curve), $\bar{\sigma}_{tCNN}$ (green points), and relative error (red triangles) given by $100(\bar{\sigma}_{tCNN} - \bar{\sigma}_{RM})/\bar{\sigma}_{RM}$ versus \bar{F}_{11} .

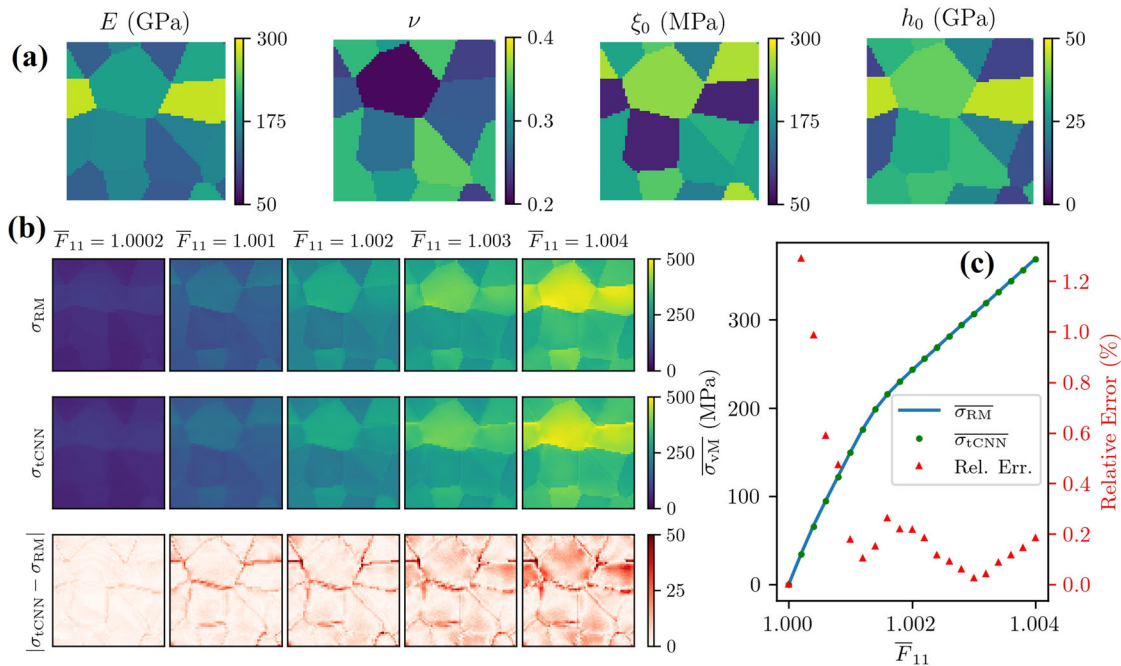


Fig. 2 Ten-grain test microstructure subject to uniaxial extension along the horizontal axis. **a** Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , linear isotropic hardening modulus h_0 . **b** Results for σ_{RM} (above), σ_{tCNN} (middle), $|\sigma_{tCNN} - \sigma_{RM}|$ (below) (all MPa), at different \bar{F}_{11} . **c** $\bar{\sigma}_{RM}$ (blue curve), $\bar{\sigma}_{tCNN}$ (green points), and relative error (red triangles) given by $100(\bar{\sigma}_{tCNN} - \bar{\sigma}_{RM})/\bar{\sigma}_{RM}$ versus \bar{F}_{11} .

Extension histories not in the training dataset

The training data is based in particular on results for σ_{RM} corresponding to uniaxial extension of the unit cell up to $\bar{F}_{11} = 1.004$. Since σ_{vM} is an input to the network, the tCNN can be used to calculate σ_{vM} for values larger than those in the training data, corresponding to unit cell uniaxial extension beyond that represented by the training data. A typical maximum is just

above 500 MPa (see, for example, Fig. 11, bottom right). Results for this are shown in Fig. 8.

Note the significant increase in $|\sigma_{tCNN} - \sigma_{RM}|$ in Fig. 8b (bottom) and in the corresponding relative error in Fig. 8c (red triangles) above $\bar{F}_{11} = 1.0065$.

Figure 8c shows the corresponding stress-strain curve for further loading. According to this curve, the error notably

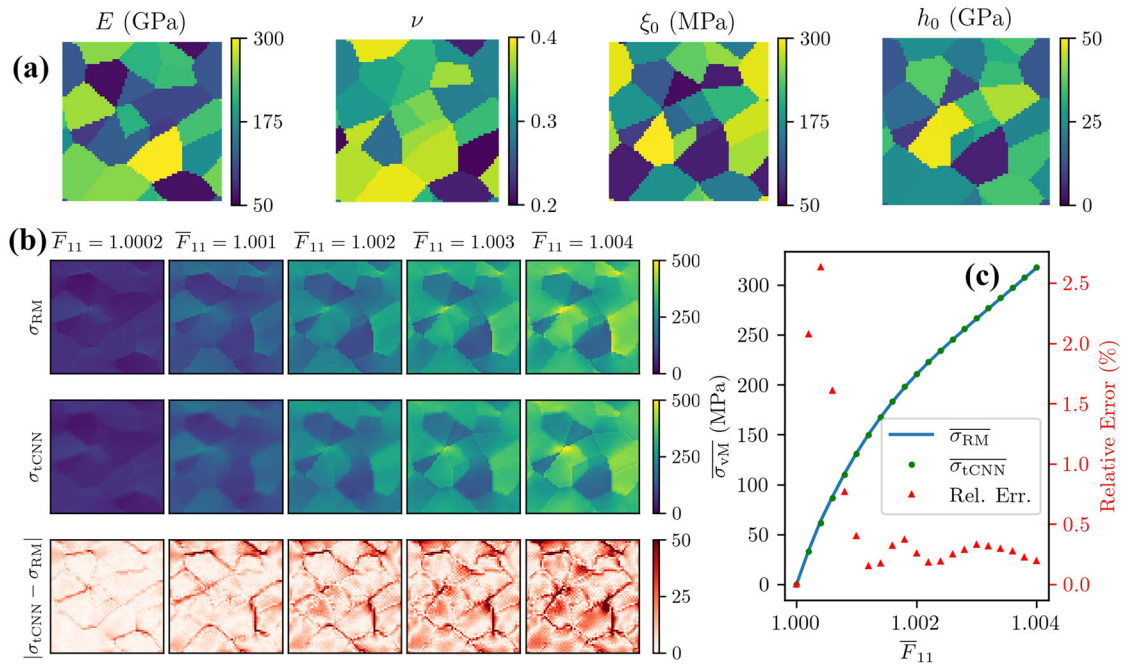


Fig. 3 Twenty-grain test microstructure subject to uniaxial extension along the horizontal axis. **a** Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , linear isotropic hardening modulus h_0 . **b** Results for σ_{RM} (above), σ_{tCNN} (middle), $|\sigma_{tCNN} - \sigma_{RM}|$ (below) (all MPa), at different \bar{F}_{11} . **c** $\bar{\sigma}_{RM}$ (blue curve), $\bar{\sigma}_{tCNN}$ (green points), and relative error (red triangles) given by $100(\bar{\sigma}_{tCNN} - \bar{\sigma}_{RM})/\bar{\sigma}_{RM}$ versus \bar{F}_{11} .

Case	1	2	3	4
MAE (MPa)	3.662	1.925	1.519	1.265

increases for values of $\bar{\sigma}_{vM}$ greater than 500 MPa. One reason for this is very little data in the training data for such values. This can be seen for example in Fig. 11 (below right), which displays the distribution of pixel values for σ_{vM} at $\bar{F}_{11} = 1.004$. As shown, only 0.084% of the pixels in the whole training dataset have a von Mises stress level above 500 MPa.

Computational efficiency

We consider next the time required to calculate the von Mises stress field based on the RM and the tCNN. Recall that the former is based on the numerical solution of BVPs for quasi-static mechanical equilibrium in periodic unit cells employing spectral numerical solution methods (e.g., refs. 1,2,4,7). The corresponding results for the RM are obtained using the DAMASK simulation package⁵ (see next section). The calculations are carried out on a single core of an Intel® Core™ i9900K clocked at 3.60 GHz. On this basis, the run time of DAMASK for a single training simulation averaged over 50 training simulations is approximately 75 s, and the corresponding run time for the tCNN is about 0.15 s. As such, the tCNN is about 500 faster than the RM for the training case. Of course, this difference depends on the details of the training data, e.g., on the chosen resolution of 64×64 pixels. For finer resolutions, the difference increases significantly.

DISCUSSION

In this work, a U-Net-based CNN has been trained to calculate the von Mises stress field in metallic polycrystals or composites in which the mechanical behavior of the grains is modeled by

J_2 elasto-viscoplasticity with linear isotropic hardening. Data for training, validation, and testing were generated via numerical solution of the corresponding periodic boundary-value problems for quasi-static mechanical equilibrium on grain microstructure unit cells via spectral methods. Data sets for testing of the resulting trained CNN (tCNN) are based on the (i) number of grains, (ii) distribution of material properties, (iii) grain morphologies, and (iv) applied unit cell extension.

For the 64×64 pixel resolution of the microstructure employed, calculation of the von Mises stress field σ_{vM} with the resulting tCNN is up to 500 times faster than with the RM. Increasing the resolution would result in an even larger difference in computational time. On the other hand, in contrast to the RM, the accuracy of the tCNN is limited to (i) the range of the training data as well as (ii) uniaxial extension.

The current approach can be extended and further developed in a number of directions. These include for example (i) output of all components of the stress field, (ii) training for different deformation and loading conditions, or (iii) training for multiple rates of deformation and/or loading. As well, training can be based on more sophisticated “physics-informed” loss functions (accounting, for example, for mechanical equilibrium), or on more sophisticated “physics-encoded” network architectures. This represents work in progress to be reported on in the future.

METHODS

RM: J_2 viscoplasticity with isotropic hardening

As discussed in the introduction, data for the training and testing of the ANN are obtained from the numerical solution of a boundary-value problem for quasi-static mechanical equilibrium based on viscoplastic material modeling of grain behavior in a grain microstructure. In the current context of isothermal and quasi-static conditions, these include in particular mechanical equilibrium $\text{div} \mathbf{P} = \mathbf{0}$ in terms of the first Piola–Kirchhoff stress \mathbf{P} . In the context of the viscoplastic decomposition $\mathbf{F} = \mathbf{F}_e \mathbf{F}_p$ of the deformation gradient \mathbf{F} , the linear elastic relation $\mathbf{S}_e = \lambda(\mathbf{I} \cdot \mathbf{E}_e) + 2\mu \mathbf{E}_e \cdot \mathbf{E}_e$ is

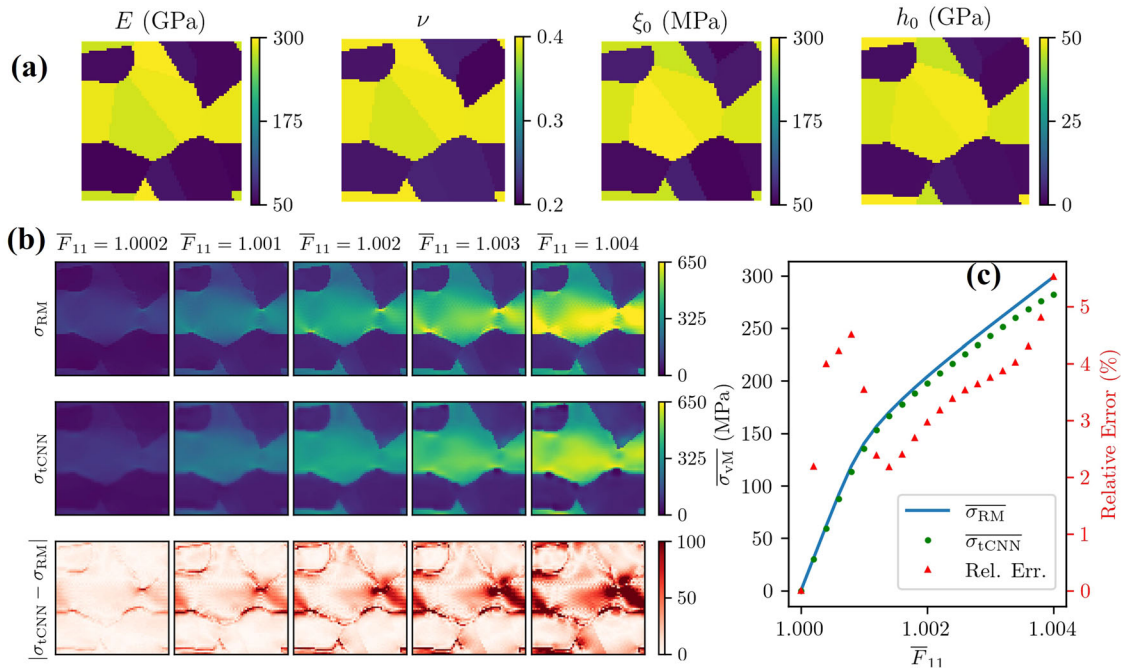


Fig. 4 Case 1 (high-contrast) test microstructure subject to uniaxial extension along the horizontal axis. a Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , linear isotropic hardening modulus h_0 . **b** Results for σ_{RM} (above), σ_{tCNN} (middle), $|\sigma_{tCNN} - \sigma_{RM}|$ (below) (all MPa), at different \bar{F}_{11} . **c** $\bar{\sigma}_{RM}$ (blue curve), $\bar{\sigma}_{tCNN}$ (green points), and relative error (red triangles) given by $100(\bar{\sigma}_{tCNN} - \bar{\sigma}_{RM})/\bar{\sigma}_{RM}$ versus \bar{F}_{11} .

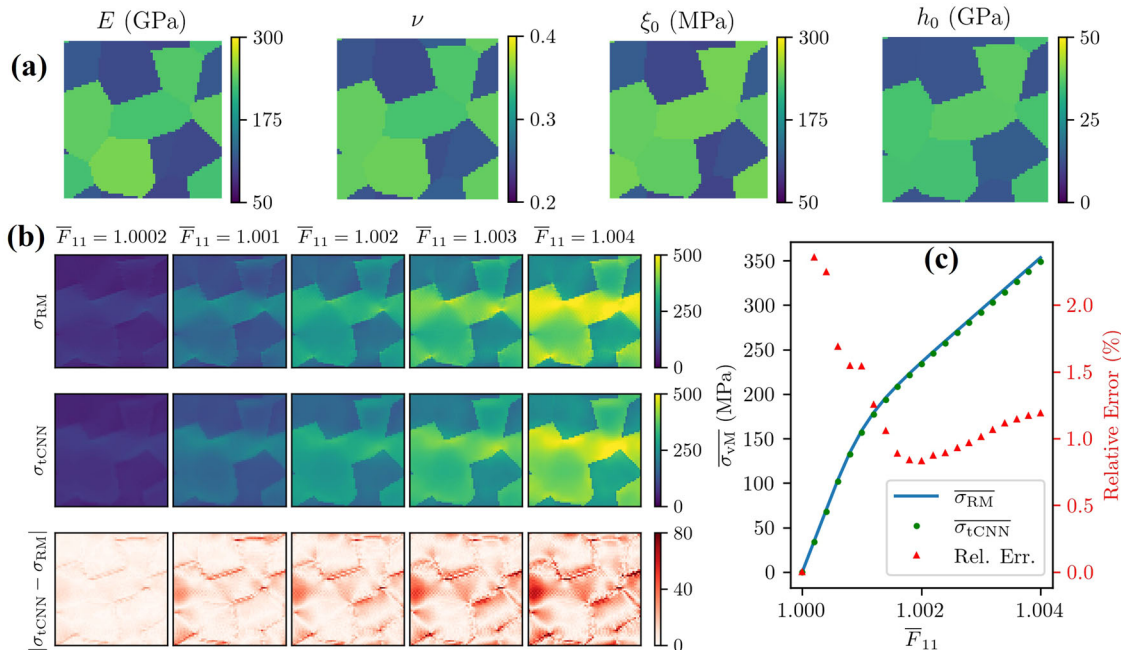


Fig. 5 Case 3 (low-contrast) test microstructure subject to uniaxial extension along the horizontal axis. a Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , linear isotropic hardening modulus h_0 . **b** Results for σ_{RM} (above), σ_{tCNN} (middle), $|\sigma_{tCNN} - \sigma_{RM}|$ (below) (all MPa), at different \bar{F}_{11} . **c** $\bar{\sigma}_{RM}$ (blue curve), $\bar{\sigma}_{tCNN}$ (green points), and relative error (red triangles) given by $100(\bar{\sigma}_{tCNN} - \bar{\sigma}_{RM})/\bar{\sigma}_{RM}$ versus \bar{F}_{11} .

Table 3. Material property ranges used for generating the dataset.

	E (GPa)	ν	ξ_0 (MPa)	h_0 (GPa)
Minimum	50	0.2	50	0
Maximum	300	0.4	300	50

assumed between the elastic second Piola-Kirchhoff stress \mathbf{S} and the elastic Green strain $\mathbf{E}_e = \frac{1}{2}(\mathbf{F}_e^T \mathbf{F}_e - \mathbf{I})$; note that $\mathbf{P} = \mathbf{F}_e \mathbf{S}_e \mathbf{F}_p^T$. In the following, the Lamé constants with $\lambda = E\nu/((1+\nu)(1-2\nu))$ and $\mu = E/(2(1+\nu))$ are determined in terms of the Young's modulus E and the Poisson ratio ν . The evolution of \mathbf{F}_p

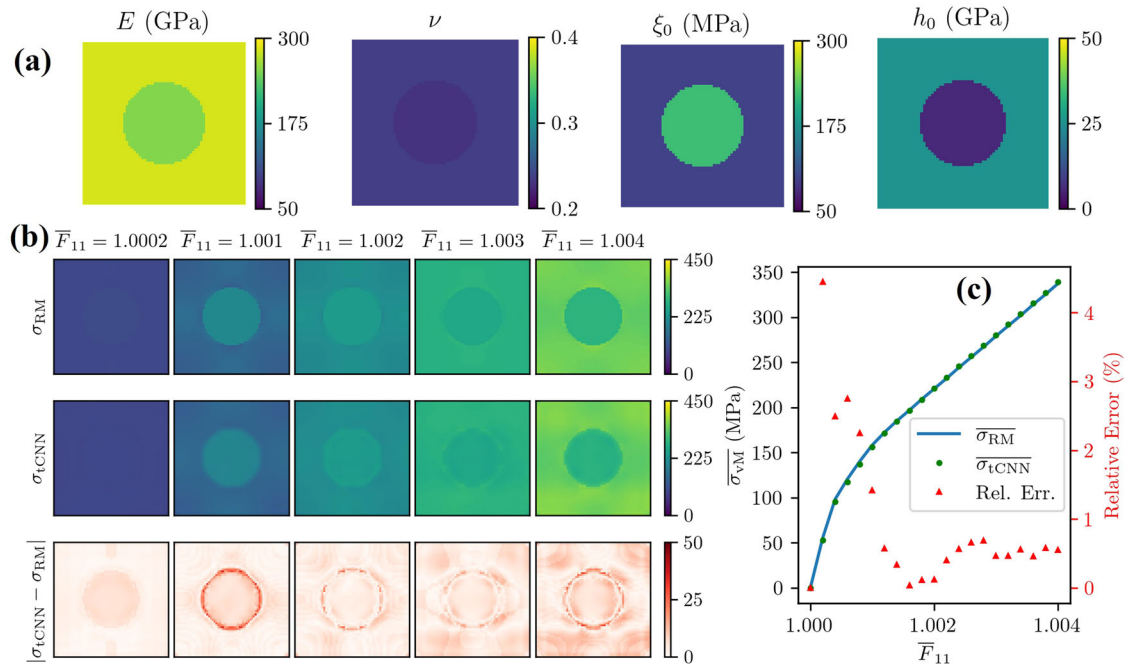


Fig. 6 Round inclusion test microstructure subject to uniaxial extension along the horizontal axis. a Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , linear isotropic hardening modulus h_0 . **b** Results for σ_{RM} (above), σ_{tCNN} (middle), $|\sigma_{tCNN} - \sigma_{RM}|$ (below) (all MPa), at different \bar{F}_{11} . **c** σ_{vM} (blue curve), σ_{tCNN} (green points), and relative error (red triangles) given by $100(\sigma_{tCNN} - \sigma_{RM})/\sigma_{RM}$ versus \bar{F}_{11} .

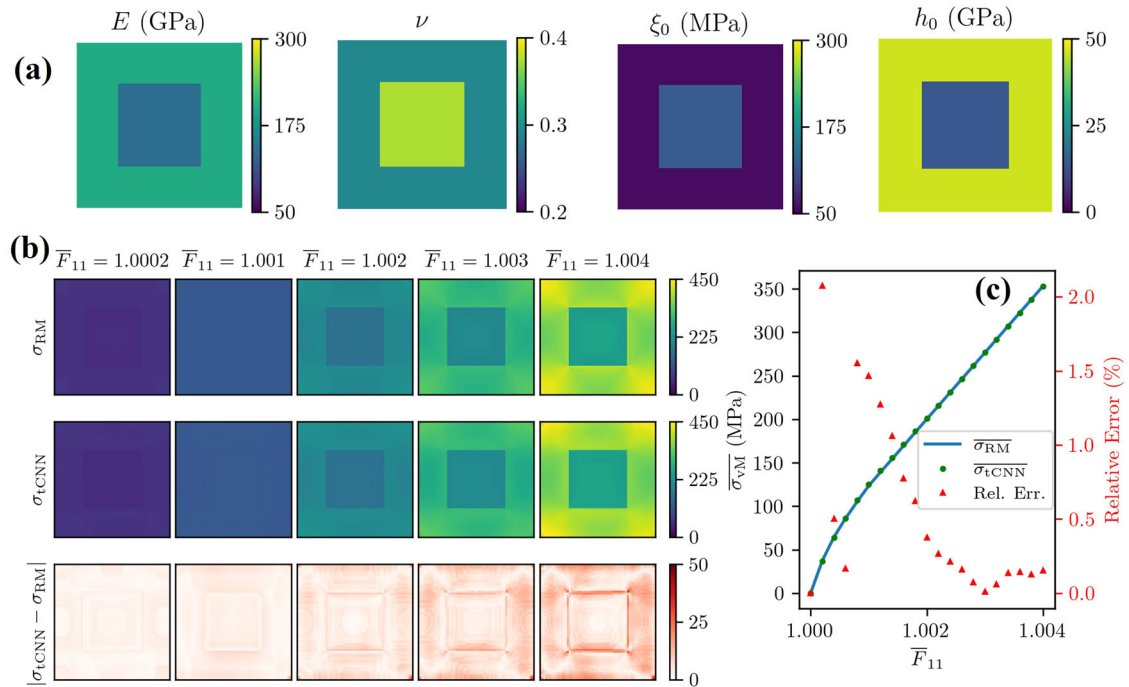


Fig. 7 Square inclusion test microstructure subject to uniaxial extension along the horizontal axis. a Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , linear isotropic hardening modulus h_0 . **b** Results for σ_{RM} (above), σ_{tCNN} (middle), $|\sigma_{tCNN} - \sigma_{RM}|$ (below) (all MPa), at different \bar{F}_{11} . **c** σ_{vM} (blue curve), σ_{tCNN} (green points), and relative error (red triangles) given by $100(\sigma_{tCNN} - \sigma_{RM})/\sigma_{RM}$ versus \bar{F}_{11} .

is determined by the J_2 flow rule $\dot{\mathbf{F}}_p \mathbf{F}_p^{-1} = \dot{\gamma}_p \mathbf{S}_e^{\text{dev}} / |\mathbf{S}_e^{\text{dev}}|$, where $\dot{\gamma}_p$ is the rate of equivalent plastic shear, and $\mathbf{S}_e^{\text{dev}}$ the deviatoric part of \mathbf{S}_e . The viscoplastic (i.e., rate-dependent²⁸) form $\dot{\gamma}_p = \dot{\gamma}_0 (|\mathbf{S}_e^{\text{dev}}|/\xi)^{n_0}$ for the evolution of γ_p is determined by the typical material inelastic shear rate $\dot{\gamma}_0$, the power-law exponent n_0 , and the flow resistance $\xi(\gamma_p) = \xi_0 + h_0 \gamma_p$ for linear isotropic

hardening, with ξ_0 the initial flow resistance, and h_0 the isotropic hardening modulus.

The viscoplastic model is implemented in the simulation software toolkit DAMASK⁵. This toolkit is also used for numerical solution of the corresponding quasi-static mechanical boundary-value problem on periodic grain microstructures based on spectral (i.e., Fourier) methods.

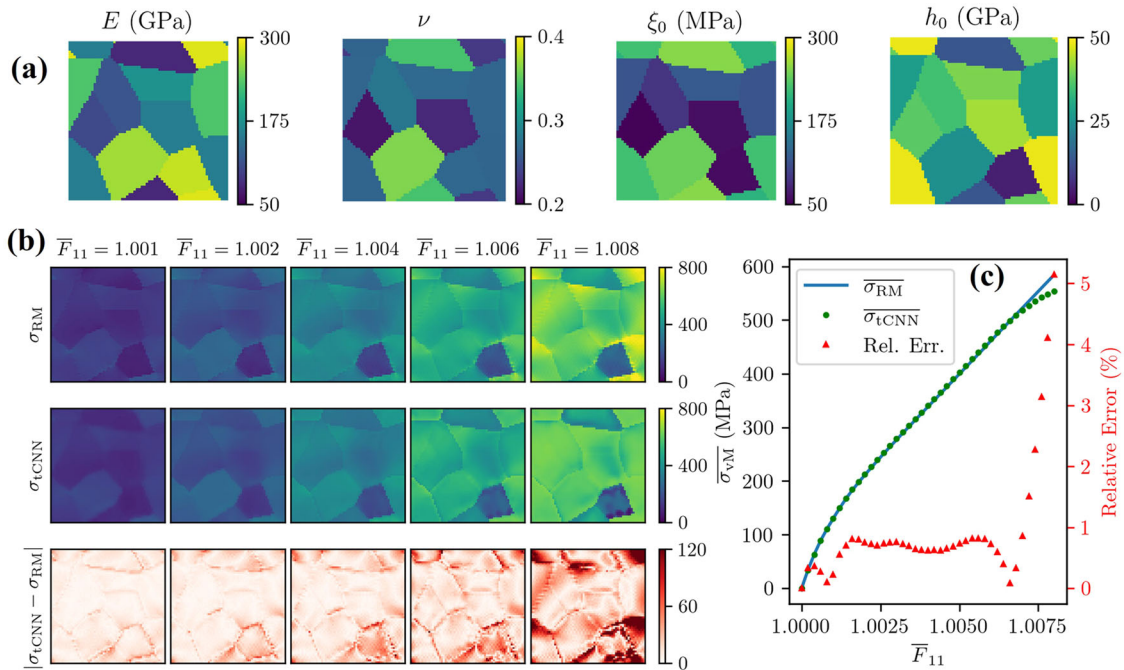


Fig. 8 Extension histories test microstructure subject to uniaxial extension along the horizontal axis. **a** Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , linear isotropic hardening modulus h_0 . **b** Results for σ_{RM} (above), σ_{tCNN} (middle), $|\sigma_{tCNN} - \sigma_{RM}|$ (below) (all MPa), at different \bar{F}_{11} . **c** $\bar{\sigma}_{RM}$ (blue curve), $\bar{\sigma}_{tCNN}$ (green points), and relative error (red triangles) given by $100(\bar{\sigma}_{tCNN} - \bar{\sigma}_{RM})/\bar{\sigma}_{RM}$ versus \bar{F}_{11} .

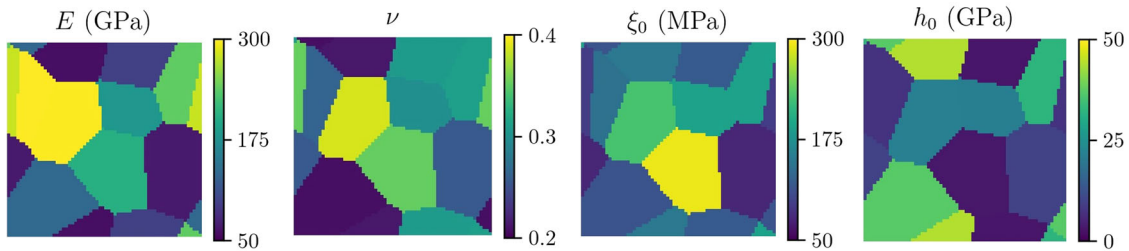


Fig. 9 Example of a randomly generated grain microstructure and corresponding material property distribution. The latter include Young's modulus E , Poisson's ratio ν , the initial flow resistance ξ_0 , and the linear isotropic hardening modulus h_0 .

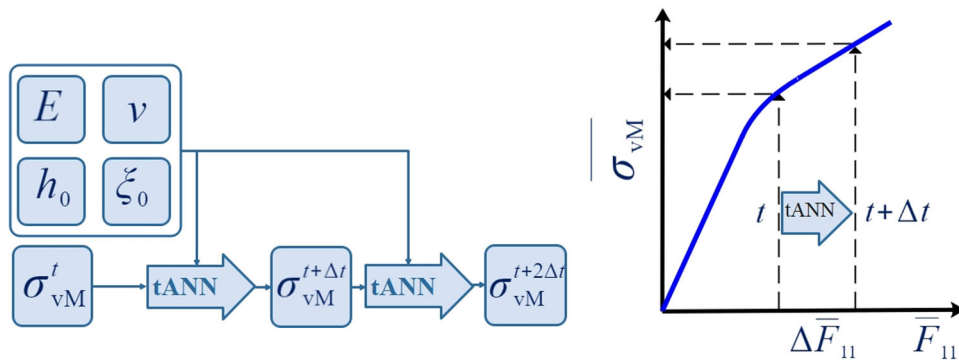


Fig. 10 Schematic illustration of the tANN for calculation of the local von Mises stress based on J_2 viscoplasticity with linear isotropic hardening. **a** Young's modulus E , Poisson's ratio ν , initial flow resistance ξ_0 , and linear isotropic hardening modulus h_0 are material property distributions as input of the tANN. σ_{vM}^t denotes the von Mises stress distribution at time step t .

Data generation

Material parameters for each grain in the microstructure include E , ν , ξ_0 , h_0 , $\dot{\gamma}_0$, and n_0 . For simplicity, $\dot{\gamma}_0 = 10^{-3} \text{ s}^{-1}$ and $n_0 = 20$ are assumed the same for all grains. Values for the remaining material properties of each grain are chosen randomly from a range of

values shown in Table 3. The training dataset consists of 1000 grain microstructures based on 10 grains and random material property distributions. Grain morphologies and microstructures are generated randomly via Voronoi tessellation. An example is shown in Fig. 9.

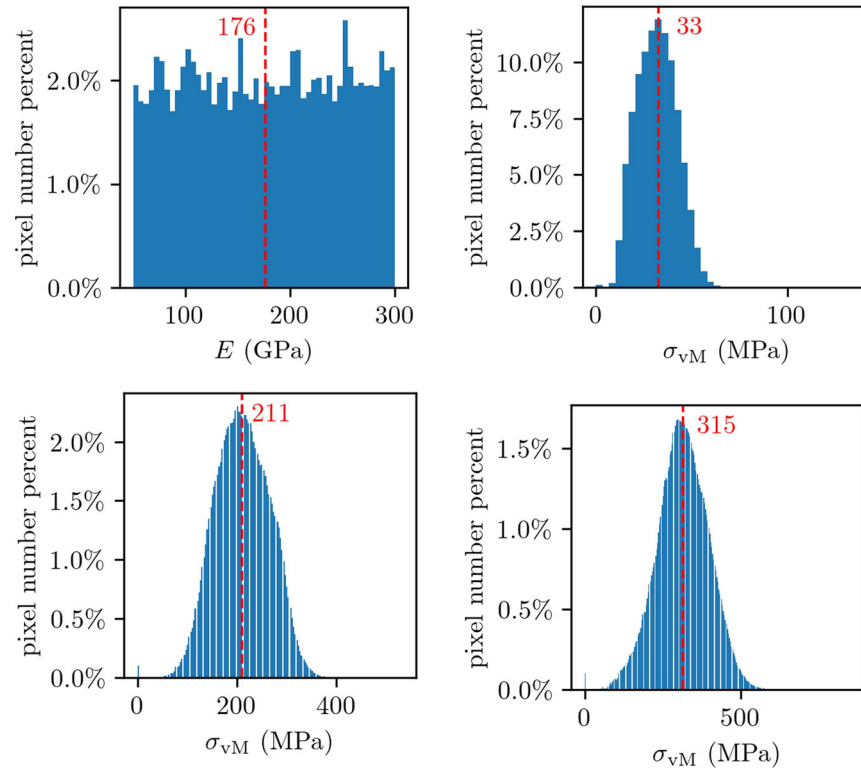


Fig. 11 Distributions of values and corresponding mean value (in red) for selected input quantities in the training data. Top left: Young's modulus E . Top right: σ_{vM} for $\bar{F}_{11} = 1.001$. Bottom left: σ_{vM} for $\bar{F}_{11} = 1.002$. Bottom right: σ_{vM} for $\bar{F}_{11} = 1.004$.

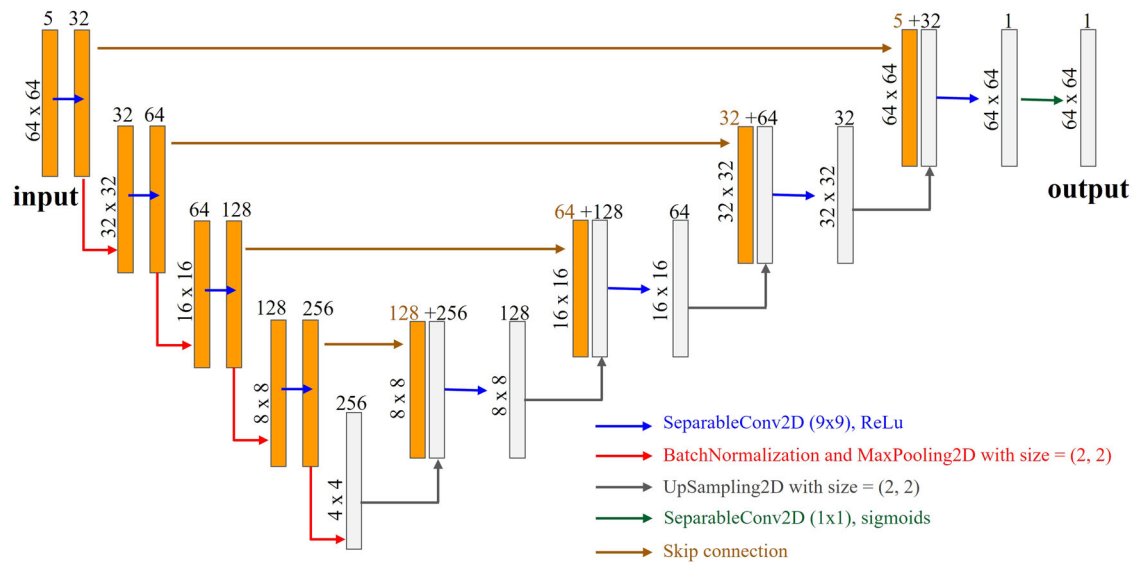


Fig. 12 U-Net-based ANN architecture adopted in this work. Depicted are the number of channels as well as the size of the images in each layer. Employed are two-dimensional (2D) separable convolution with 9×9 kernel, ReLU activation, batch normalization, 2D max pooling, and upsampling with bilinear interpolation, all as implemented in TensorFlow³⁰. All weights and biases are initialized via the Glorot algorithm³¹. See text for discussion.

As done, for example, in ref. ²³, the ANN input consists of the distribution of these material properties in the unit cell/microstructure as well as results for the scalar von Mises stress field $\sigma_{vM} = \sqrt{3\mathbf{T}_{dev} \cdot \mathbf{T}_{dev}/2}$ (Cauchy stress $\mathbf{T} = \mathbf{P}\mathbf{F}^T/\det\mathbf{F}$) during extension. Given the material property distribution (E, ν, h_0, ξ_0) and σ_{vM} at time step t (i.e. σ_{vM}^t), then, the trained

ANN (tANN) outputs σ_{vM} at time step $t + \Delta t$ (i.e., $\sigma_{vM}^{t+\Delta t}$). This is depicted in Fig. 10.

In the current case of purely bulk behavior of grain microstructures on the unit cell U , fields $f = \bar{f} + \tilde{f}$ on the unit cell U are additively split into mean \bar{f} and fluctuation \tilde{f} parts, with $\bar{f} := v(U)^{-1} \int_U f dv$ and $v(U) := \int_U dv$. In this context, deformation

“boundary conditions” on U take the form of prescribed values for the mean deformation gradient $\bar{\mathbf{F}}(t)$. For the current case of uniaxial extension, the Cartesian/matrix form

$$\bar{\mathbf{F}}(t) \equiv \begin{bmatrix} \bar{F}_{11}(t) & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \bar{F}_{11}(t) = 1 + \dot{\bar{F}}_{11}(0) t, \quad (2)$$

of $\bar{\mathbf{F}}(t)$ applies. For simplicity, data for training and testing are limited to a single extension rate $\dot{\bar{F}}_{11}(0) = 1 \times 10^{-3} \text{ s}^{-1}$, i.e., to quasi-static extension. Based on this extension rate, results obtained up to $\bar{F}_{11}(4) = 1.004$ are employed as training data, and those between this value and $\bar{F}_{11}(8) = 1.008$ are used for testing the tCNN.

The pixel distributions of values for Young’s modulus E and σ_{VM} for different values of \bar{F}_{11} in the training data are shown in Fig. 11.

The pixel distributions of the other input material properties are similar. As evident, the pixel distribution of σ_{VM} at each value of \bar{F}_{11} is quasi-normal in character, as expected for randomly distributed material property values.

Neural network type, architecture, and training

The current ANN is based on the U-Net convolutional type and architecture introduced by Ronneberger et al.²⁹. As shown by Mianroodi et al.²³, this architecture is suitable for surrogate modeling the stress field in solid mechanics problems, and in particular the von Mises stress σ_{VM} . Fig. 12 schematically depicts the U-Net network architecture, (referred to in ref. ²⁹ as U-Net due to its shape).

Network input consists of 64×64 pixel images for E , ν , ξ_0 , h_0 , and σ_{VM} , and the network outputs one 64×64 pixel image for σ_{VM} . As usual, both input and output are normalized to 1. In addition, 32 filters capture the main features from the input images. As shown in the figure, four types of operation are performed in the U-Net, namely, separable two-dimensional (2D) convolution with a kernel size of 9, batch normalization, 2D max pooling, and 2D upsampling with bilinear interpolation.

For the training process, Adam optimization is employed with a learning rate of 0.001 and a momentum of 0.9. As discussed above, the loss function for training, validation, and testing is given by the mean absolute error (MAE) based on the difference $\sigma_{\text{tCNN}} - \sigma_{\text{RM}}$. TensorFlow³⁰ is used to set up and train the network. As usual, the dataset is divided into training (80%) and (20%) testing subsets. Training is based on 500 epochs and has an MAE of 1.733 MPa; the corresponding MAE for the validation dataset is 1.743 MPa. No sign of overfitting was observed.

DATA AVAILABILITY

The data employed in the current work can be obtained from the corresponding author upon request.

CODE AVAILABILITY

The code employed in the current work can be obtained from the corresponding author upon request.

Received: 2 September 2022; Accepted: 21 February 2023;

Published online: 13 March 2023

REFERENCES

- Shanthraj, P., Eisenlohr, P., Diehl, M. & Roters, F. Numerically robust spectral methods for crystal plasticity simulations of heterogeneous materials. *Int. J. Plast.* **66**, 31–45 (2015).
- Willot, F. Fourier-based schemes for computing the mechanical response of composites with accurate local fields. *C. R. Mécanique* **343**, 232–245 (2015).

- Schneider, M., Merkert, D. & Kabel, M. FFT-based homogenization for microstructures discretized by linear hexahedral elements. *Int. J. Numerical Methods Eng.* **109**, 1461–1489 (2017).
- Lucarini, S. & Segurado, J. DBFFT: a displacement based FFT approach for non-linear homogenization of the mechanical behavior. *Int. J. Eng. Sci.* **144**, 103131 (2019).
- 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).
- Lebensohn, R. A. & Rollett, A. D. Spectral methods for full-field micromechanical modelling of polycrystalline materials. *Comput. Mater. Sci.* **173**, 109336 (2020).
- Khorrami, M., Mianroodi, J. R., Shanthraj, P. & Svendsen, B. Development and comparison of spectral algorithms for numerical modeling of the quasi-static mechanical behavior of inhomogeneous materials. Preprint at *arXiv* <https://arxiv.org/abs/2009.03762> (2020).
- 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).
- Roters, F. et al. DAMASK: the Düsseldorf Advanced Material Simulation Kit for studying crystal plasticity using an FE based or a spectral numerical solver. *Proc. IUTAM* **3**, 3–10 (2012).
- Diehl, M. et al. Identifying structure–property relationships through DREAM. 3D representative volume elements and DAMASK crystal plasticity simulations: an integrated computational materials engineering approach. *JOM* **69**, 848–855 (2017).
- Stoll, A. & Benner, P. Machine learning for material characterization with an application for predicting mechanical properties. *GAMM-Mitteilungen* **44**, e202100003 (2021).
- Wu, X. *Neural Network-Based Material Modeling*. PhD thesis, University of Illinois at Urbana-Champaign (1991).
- Haj-Ali, R., Kim, H.-K., Koh, S. W., Saxena, A. & Tummala, R. Nonlinear constitutive models from nanoindentation tests using artificial neural networks. *Int. J. Plast.* **24**, 371–396 (2008).
- Ali, U., Muhammad, W., Brahme, A., Skiba, O. & Inal, K. Application of artificial neural networks in micromechanics for polycrystalline metals. *Int. J. Plast.* **120**, 205–219 (2019).
- Mayer, A. E., Krasnikov, V. S. & Pogorelko, V. V. Dislocation nucleation in Al single crystal at shear parallel to (111) plane: molecular dynamics simulations and nucleation theory with artificial neural networks. *Int. J. Plast.* **139**, 102953 (2021).
- Pandya, K. S., Roth, C. C. & Mohr, D. Strain rate and temperature dependent fracture of aluminum alloy 7075: experiments and neural network modeling. *Int. J. Plast.* **135**, 102788 (2020).
- Settgast, C., Hütter, 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).
- Mianroodi, J. R., Rezaei, S., Siboni, N. H., Xu, B.-X. & Raabe, D. Lossless multi-scale constitutive elastic relations with artificial intelligence. *npj Comput. Mater.* **8**, 67 (2022).
- Raissi, M., Perdikaris, P. & Karniadakis, G. E. 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).
- Lagaris, I. E., Likas, A. & Fotiadis, D. I. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Trans. Neural Netw.* **9**, 987–1000 (1998).
- 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).
- 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).
- Mianroodi, J. R., Siboni, N. & Raabe, D. Teaching solid mechanics to artificial intelligence—a fast solver for heterogeneous materials. *npj Comput. Mater.* **7**, 1–10 (2021).
- Rashidi, M. M., Pittie, T., Chakraborty, S. & Krishnan, N. M. A. Learning the stress-strain fields in digital composites using Fourier neural operator. *IScience* **25**, 105452 (2022).
- Li, Z. et al. Fourier neural operator for parametric partial differential equations. Preprint at *arXiv* <https://arxiv.org/abs/2010.08895> (2020).
- Kovachki, N., Lanthaler, S. & Mishra, S. On universal approximation and error bounds for Fourier neural operators. *J. Mach. Learn. Res.* **22**, 1–76 (2021).
- Kapoor, S., Mianroodi, J. R., Khorrami, M., Siboni, N. S. & Svendsen, B. Comparison of two artificial neural networks trained for the surrogate modeling of stress in materially heterogeneous elastoplastic solids. Preprint at *arXiv* <https://arxiv.org/abs/2210.16994> (2022).
- Peirce, D., Asaro, R. J. & Needleman, A. Material rate dependence and localized deformation in crystalline solids. *Acta Metall.* **31**, 1951–1976 (1983).
- Ronneberger, O., Fischer, P. & Brox, T. U-net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention* 234–241 (Springer, 2015).

30. Abadi, M. et al. TensorFlow: large-scale machine learning on heterogeneous systems. Software available. <https://www.tensorflow.org/> (2015).
31. X., Glorot, X. & Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In *Proc. Thirteenth International Conference on Artificial Intelligence and Statistics* 249–256 (MLR, 2010).

ACKNOWLEDGEMENTS

Financial support for the current work was provided by BiGmax (<https://www.bigmax.mpg.de/>), the Max Planck research network on big-data-driven materials science.

AUTHOR CONTRIBUTIONS

M.S.K. and J.R.M. developed the initial concept and model formulation. M.S.K. carried out the solid mechanics simulations to generate the training and test datasets. N.H.S. designed the neural network, and M.S.K. obtained the neural network parameters. M.S.K., J.R.M., P.G., and B.S. wrote the initial draft. B.S., P.B., and D.R. supervised the work and contributed to the interpretation of the results.

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 Mohammad S. Khorrami or Jaber R. 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) 2023